JOURNAL OF APPLIED BEHAVIOR ANALYSIS 1989, 22, 181–196 NUMBER 2 (SUMMER 1989)

ACQUISITION, GENERALIZATION, AND MAINTENANCE OF QUESTION-ANSWERING SKILLS IN AUTISTIC CHILDREN

KRISTIN E. SECAN

UNIVERSITY OF MARYLAND AND MONTGOMERY COUNTY PUBLIC SCHOOLS

ANDREW L. EGEL

UNIVERSITY OF MARYLAND

CYNTHIA S. TILLEY

UNIVERSITY OF MARYLAND AND MONTGOMERY COUNTY PUBLIC SCHOOLS

We conducted an investigation to evaluate the effects of a training strategy for teaching autistic students generalized responses to three forms of wh— questions (what, how, and why). Students were taught, using modeling and reinforcement procedures, to answer questions with magazine pictures as the referents. Each question form was divided into two or more subcomponents reflective of common social usage and was taught within the context of a modified multiple probe design across subcomponents. Following acquisition of each subcomponent, generalization to natural context and storybook questions was assessed; additional probes were conducted to assess responding over time and whether acquisition of responses to questions promoted question-asking skills. Results showed that the picture training procedure was effective in teaching a generalized response to questions for which the relevant cue was visible, whereas specific generalization programming was required for situations in which the relevant cue was not visible. All acquired responses were durable over time.

DESCRIPTORS: question-answering, generalization, language, autistic children, developmentally disabled children

One of the most prominent characteristics of children labeled autistic or autistic-like is a severe impairment or delay in their ability to produce and respond to language (e.g., Alpert & Rogers-Warren, 1985; Carr, 1982; Lovaas, 1977; Ritvo & Freeman, 1978; Schopler & Mesibov, 1986; Schuler, 1980; Wetherby, 1986). Given the recognized impact of severe language deficits on all facets of the autistic individual's life, researchers have developed and evaluated numerous strategies for effectively teaching language skills to autistic and other severely handicapped individuals (e.g., Guess, 1969; Guess, Sailor, Rutherford, & Baer, 1968; Lovaas, 1977; Sailor, 1971; Wolf, Risley, & Mees, 1964). Much of the earliest language research emphasized the acquisition of specific linguistic forms (e.g., Guess et al., 1968; Lutzker & Sherman, 1974; Sailor, 1971).

The results of research such as that reviewed above have contributed greatly to the advancement of our understanding of language acquisition. However, many educators and researchers have shifted away from elaborating grammar and structure to examine the communicative value of language skills taught to developmentally disabled students (e.g., Alpert & Rogers-Warren, 1985; Carr, 1986; Halle, 1982; Lord, 1986).

Fay and Schuler (1980) suggested that the abilities to comprehend and use varied forms of expression and the ability to participate actively in social interchanges with others represent two skills in-

This article is based on a thesis submitted by Kristin Secan to the Department of Special Education, University of Maryland. The research was supported by Grant G00-85-30232 from the USDOE/OSERS/RPB. However, the opinions expressed herein do not necessarily reflect the position or policy of the agency, and no official endorsement by the agency should be inferred.

The authors thank Kathy Gradel, Francy Kohl, and Nancy Neef for their valuable comments on an earlier draft, and Maureen Dwyer, Maribeth Sullivan, Patty Hallinan, and Jan Greenburg for their assistance in conducting this investigation.

Correspondence and requests for reprints may be sent to Andrew L. Egel, Department of Special Education, University of Maryland, College Park, Maryland 20742.

volved in acquiring communicative competency. Schreibman and Carr (1978) addressed this issue by teaching autistic children to answer "I don't know" when presented with questions to which they had no appropriate response. Such responses are likely to increase social interchanges because they provide other persons with prompts to continue the interaction, as well as naturally setting the occasion for others to provide information to the autistic individual. Neef, Walters, and Egel (1984) used an embedded instruction procedure to teach autistic children an appropriate yes/no response to mand and tact items. Such responses enable language-disabled children to engage in interactions with and provide information to others in their environment.

One feature of most social and instructional interchanges, regardless of specific content, involves the asking and answering of questions. For example, much of the informational exchange experienced by developmentally normal children occurs in question–answer situations. The verbal stimulation necessary to facilitate language development frequently involves the use of questions, with wh— question forms predominating (Nelson, 1973; Parnell & Amerman, 1984). Furthermore, there is evidence to suggest that children learn to generate particular forms of wh— questions only after they have first learned to respond to those forms (Bloom & Lahey, 1978; Brown, 1968; Hood, 1977). Paul (1985) has included question-asking and -answering skills as important to the development of appropriate conversational skills. Thus, the ability to respond to and answer questions in a conversational context has implications for language and cognitive development, in addition to being an important component of conversational skills. Given such information, the relative paucity of research on developing question-asking and -answering skills in autistic and other developmentally delayed children is surprising. To date, little attention has been paid to teaching these individuals to generate or respond to questions beyond a few specific forms (Guess et al., 1968; Lovaas, 1977; Twardosz & Baer, 1973).

Krantz, Zalenski, Hall, Fenske, and McClannahan (1981) conducted more extensive investigations of question-answering skills. The study most relevant to our investigation involved teaching autistic children to answer three classes of questions (what, how, and why) referred to as wh— concepts. Results of this investigation indicated that the procedure was effective in teaching a generalized response to three wh— concepts when magazine pictures were the referents for training and generalization questions. Such results are encouraging, in that the authors documented a simple and economical approach to teaching a complex language skill. However, Krantz et al. (1981) assessed generalization only across novel pictures; no data were obtained on whether students could answer questions in more social, communicative contexts. Furthermore, both training and generalization sessions were conducted by the same individual. Whether such training will result in broader generalization has not been established. Thus, our study was designed to assess whether question-answering skills would generalize across persons and situations, as well as maintain over time.

METHOD

Participants and Setting

Participants were 4 students (1 female, 3 males) who attended a special education program for autistic and autistic-like children. All students had been diagnosed as autistic by independent professionals in accordance with the diagnostic criteria established by the National Society for Children and Adults with Autism (Ritvo & Freeman, 1978).

All students were delayed in social interaction (rarely interacting with peers) and play skills (using toys and other leisure materials in a stereotypic or non-age-appropriate manner), exhibited a variety of stereotypies (flapping, tapping and twirling objects, verbal perseveration, finger flipping, and light gazing), and were delayed in language and communication skills. Additionally, a majority of these students required systematic monitoring of behavior problems such as aggression, noncompliance, tantrums, and property destruction.

All students had received daily instruction in the DISTAR Language I series (Engelmann & Osborn, 1976) for 1 to 5 years, depending on time of

Table 1
Student Characteristics

	Student			
	1	2	3	4
Age	9 years, 1 month	7 years, 11 months	5 years, 7 months	9 years, 2 months
IQ/assessment	40 WISC-R	66 Stanford-Binet	95 Kaufman Achievement	46 WISC-R
Language age/ assessment	3 years, 4 months: T.A.C.L.	3.10: Minnesota Child Dev. Inventory	4 years, 8 months: T.A.C.L.	4 years: T.A.C.L.
Years in DISTAR language	1	2	1	5
Social language use	Greets people spontaneously and responds to greetings. Answers personal information and what-verb questions. Spontaneously makes simple requests.	Responds to greetings, initiates wants, needs. Answers what, personal information, who, and where questions. Spontaneously labels events and objects.	Responds to greetings, initiates wants, needs. Answers what-verb questions, personal information, and who questions. Spontaneously labels objects.	Initiates greetings. Answers what, where, who questions. Initiates questions regarding daily routines.

enrollment in the school program. Additionally, all students received approximately 45 min weekly of individual and group speech and language therapy from the school speech therapy staff. Table 1 summarizes pertinent information for each student.

Stimulus Materials

Training stimuli consisted of approximately 50 pictures taken from magazines, books, and other commercially prepared materials. The photographs depicted adults and children engaged in a range of activities across a variety of domains (e.g., visiting the doctor, celebrating holidays, performing household tasks) and were selected on the basis of their relevance to the students' range of experience. Of the 50 pictures, the 15 most relevant to each different use of a given question form were selected for training; a subgroup of 10 of these was randomly selected to be used during baseline and posttraining sessions (see picture training).

Target Behavior

The target behavior was student responses to wh— questions. Wh— questions refer to both adverbial interrogative words (when, where, how, why) and nominal interrogative words (who, whom, whose, what, and which; see Krantz et al., 1981). Questions beginning with "Why," "How," and "What" (referred to as question forms) were chosen for this experiment because the students consistently demonstrated most difficulty in answering them. Students received training on the different categories or subcomponents of these question forms if they scored 60% or less during baseline probes. Therefore, some students were trained across three subcomponents of a question form; others only two. The question forms and their respective subcomponents are presented in Table 2.

Probes

Pre- and posttraining probes. Student responses to wh— questions when magazine pictures were the referents were assessed prior to (baseline) and immediately following training. The student was seated at a table facing or next to a staff member who presented a picture, directed the student to attend to the item (e.g., "Look at the picture"), and asked the corresponding wh— question. The first word in each sentence, representing the target question form, was emphasized. If a student failed to respond to the question within 10 s, the staff member moved to the next trial. No feedback was

Table 2
Targeted Question Forms and Subcomponents

Form	Example
What	
a. as an object or noun	What is in your lunch?
b. as "which"	What store is this?
Why	
a. relevant to cause/effect	Why is he wearing a coat?
b. relevant to affect	Why is she crying?
c. relevant to potential action	Why is he holding a broom?
How	
a. relating to action	How do you move the ball?
b. relating to means	How is he getting to work?
c. relating to affect	How does she feel?

given for correct or incorrect responses; however, students were praised during each intertrial interval for general on-task behaviors (e.g., attending, responding, sitting appropriately, etc.). Probe stimuli consisted of a set of 10 pictures representative of the question form and subcomponent being trained. A staff member not involved in training conducted all pre- and posttraining probes to assess generalization across persons.

Generalization probes. Generalization probes were conducted for each subcomponent prior to training and when correct responding to posttraining probes was 80% or better for two consecutive sessions. These probes were conducted by individuals not involved in any other aspect of the study. Generalization probes were terminated when students responded correctly to 80% or more of the questions for two consecutive sessions or when data stabilized at levels below this criterion. The contingencies were the same as during the pre- and posttraining probes, and the staff member emphasized the first word in each sentence. Four types of generalization probes were conducted:

1. Storybook questions: All students attended weekly library sessions, were often read to by their parents, and were involved in some type of reading program as part of their individualized education program. For these reasons, student responses to wh— questions were assessed within the context of stories. One to three short, illustrated children's stories were selected on the basis of the wh— concept being trained, the student's age, and the student's demonstrated preferences (if any). The number of books selected varied because, for some student/subcomponents, a single book did not have enough relevant pictures to obtain 10 trials.

Ten questions for each subcomponent were selected to correspond to the illustrations in the books. A student was seated with the staff member, who read the story out loud. As the staff member reached each targeted illustration, she directed the student to look at the picture and then delivered the corresponding question. Both the stories and questions were held constant, for each student, across experimental conditions.

2. Natural-context questions: This probe was designed to determine whether the training procedures were effective in promoting appropriate responses to wh— questions that might occur naturally throughout a student's day. Ten questions for each subcomponent were selected on the basis of their relevance to the student's daily routines. Whenever possible, these questions were delivered throughout the day, as appropriate opportunities occurred. For example, a student being trained to answer "How" (means subcomponent) was asked "How did you get to school?" as soon as he or she got off the school bus, with the bus still in sight. In other situations, when the referent for the question was a low-frequency event such as going to

the library, the staff member would say, "Come with me to the library. I have to return a book," and then ask, "How will we get to the library?" The questions that comprised these probes varied across students and subcomponents; however, they were constant across conditions for each student and subcomponent being trained.

3. Spontaneous questions: The third set of generalization probes was designed to assess whether students increased their frequency of spontaneous question asking as a function of training. A spontaneous question was defined as an unprompted request for information prefaced by one of the interrogatives what, how, or why. Data were collected for 5 min each morning, during arrival time. An observer was seated within 5 ft of the student and recorded the number of wh— questions asked within five consecutive 1-min intervals. These data were collected five times prior to training and five times following acquisition of each subcomponent. No contingencies were provided for question asking. The teacher did, however, answer any question addressed to her.

4. Maintenance probes: Maintenance probes on training items were conducted with the same stimuli used in pre- and posttraining probes. One probe was administered for each subcomponent using the procedures described for pre- and posttraining probes. Each time a student mastered a subcomponent, maintenance data were collected on the previously acquired subcomponents of that question form. This procedure continued across subcomponents until a student had mastered all targeted subcomponents of a given question form. A maintenance probe on the final subcomponent in each tier was conducted at the termination of the study. Thus, due to varying acquisition rates across students and question forms, the time elapsed between probes varied for each student.

Additional maintenance probes were conducted for each student to assess maintenance of responses to natural-context questions. One probe, consisting of the same questions used during natural-context generalization probes, was conducted for each subcomponent. These probes were conducted at the termination of the investigation and followed the procedures described previously for conducting natural-context generalization probes. For both sets of maintenance probes, no systematic training occurred from the time that students met the acquisition criteria to the implementation of maintenance probes.

Training

Picture training. Picture training began following stable responding during baseline probes. Each subcomponent of a question form was taught individually, in sequential fashion, to maximize the student's acquisition of the variety of responses possible to a single question form (see Table 2). Training sessions were conducted daily for 10 to 15 min and consisted of presentation of 15 different pictures and their corresponding wh— question. Ten of the pictures were those used during baseline probes, with an additional five pictures included to ensure presentation of multiple exemplars (see Stokes & Baer, 1977).

Training sessions were conducted in the same general manner as the probe sessions described above, with the critical difference being the contingencies provided for correct and incorrect responses. Correct responses were praised, and incorrect responses or failure to respond within 10 s were followed by a remedial trial, during which the teacher modeled the correct response and restated the question. Correct responses to remedial trials were praised. If the students responded incorrectly to the remedial trial, the teacher repeated the question, immediately modeled the correct response, and terminated that trial. When students responded correctly to 80% or more of the training items for two consecutive sessions, posttraining probes were administered. Generalization probes were initiated when correct responding on posttraining probes was at 80% or better for two consecutive sessions.

There were two exceptions to this procedure. Students 2 and 4 had just finished training on "how" (affect) and "why" (cause and effect), respectively, at the termination of the school year. Generalization probes were not obtainable until the onset of the following school year. At this time, posttraining and generalization probes were read-

ministered and, based on the data, picture training was reintroduced.

Remedial interspersal training. Remedial interspersal training sessions were conducted when students failed to meet criteria (two consecutive sessions, 80% or better) on probes for storybook and/or natural-context questions. If the student failed to demonstrate criterion-level generalization to either following picture training he or she received interspersal training for the item on which the lowest score was obtained. Interspersal training was conducted as follows:

1. Storybook interspersal: Ten of the picture-training stimuli for the given subcomponent were interspersed with five illustrations from a storybook. The specific questions and procedures used in interspersal training did not differ from those in the picture-training condition with the exception that every third training item was a story illustration rather than a magazine picture. However, the interspersal storybook questions were not the same as those used during storybook generalization probes.

2. Natural-context interspersal: The interspersal procedures for natural-context questions were identical to those in the picture-training condition with the exception that every third question was a natural-context question appropriate to the subcomponent being trained. These questions differed from those in the generalization probes.

The criterion for mastery of interspersal training was 80% correct responding over two consecutive sessions. When students met criterion, generalization probes were readministered for that set of generalization items. Once students met criterion on all previously failed items, training on that subcomponent was discontinued. If the interspersal training was successful in promoting generalization to those probe items represented in the training (e.g., storybook) but not to the other class of generalization items (e.g., natural context), remedial interspersal training was begun on the second class.

Direct training. In two cases, storybook interspersal training procedures did not result in responding at the criterion level on generalization probes. As a result, a direct-training procedure was implemented. In this condition, 10 questions (differing from those used in generalization probes) were selected to correspond to illustrations in the same storybook used in probes. Utilizing the same procedures as in previous training conditions, students were taught to answer storybook questions until they reached the previously stated criterion, at which time generalization probes were readministered.

Data Collection and Reliability

Correct responses were defined as those occurring within 10 s and that were reasonable and socially acceptable given the context of the picture or setting and the specific type of question presented (see Krantz et al., 1981). Such a global definition was necessary because of the variety of appropriate responses possible for each question. For example, appropriate responses to the question "Why are you wearing a coat?" (cause and effect) might include "Because I'm cold"; "Because my teacher told me to"; and "Because it is time to go out." Observers used the following rules to determine whether responses met criteria:

1. In answering questions that required labels (nouns and verbs) as part of a response, students had to use correct labels for items or actions depicted in the picture or setting.

2. To be considered reasonable, the response had to be directly related to events depicted in the picture, story, or setting. For example, if during a natural-context probe, a student was asked "Why is the teacher cleaning the floor?" and the student responded "Because there's food on it," the response was scored correct only if there was food visible on the floor or food had just been consumed in that immediate area.

3. Responses had to be appropriate to the specific question form and subcomponent being assessed. For example, when asked "Why are you going to the cafeteria?" a student response of "Because I'm hungry" would be scored correct if asked in the context of a cause and effect subcomponent but incorrect if in the context of potential action.

4. Complete sentences were not required, as long as the response contained adequate information for the questioner to ascertain that the student was, in fact, responding to the given question form and subcomponent. For example, although "Toothbrush" was scored as a correct response to "What's in the rack?" the same response would be scored incorrect if given in response to the question "How do you brush your teeth?" In this instance, the minimal correct response would be "With a toothbrush."

Interobserver reliability scores were obtained for 33% of the total sessions. Reliability observers were trained, prior to the collection of data, in the definition of each wh— question and subcomponent and the previously stated definition of correct and incorrect responses. Two graduate students in special education and a classroom instructional assistant served as reliability observers.

Reliability was calculated for question answering by dividing the total number of agreements (each observer recording the same response as correct or incorrect) by the total number of agreements plus disagreements and multiplying by 100. This formula was used to calculate agreement percentages for occurrence and nonoccurrence for each condition. The average reliability scores for occurrence were 99.5% (range, 89% to 100%), 93.2% (range, 0% to 100%), 99.4% (range, 80% to 100%), 98.3% (range, 80% to 100%), and 100% for baseline, posttraining, storybook, natural-context, and question-asking probes, respectively. Average scores for nonoccurrence were similar: 95% (range, 80% to 100%), 96.9% (range, 50% to 100%), 100%, 99.5% (range, 80% to 100%), and 100% for the same conditions listed above. The range in reliability scores reflects sessions in which students produced only one or two responses. No instances of question asking were observed by either observer.

Experimental Design

This study employed a modified multiple probe design (Horner & Baer, 1978) with replications across subcomponents, question forms, and students.

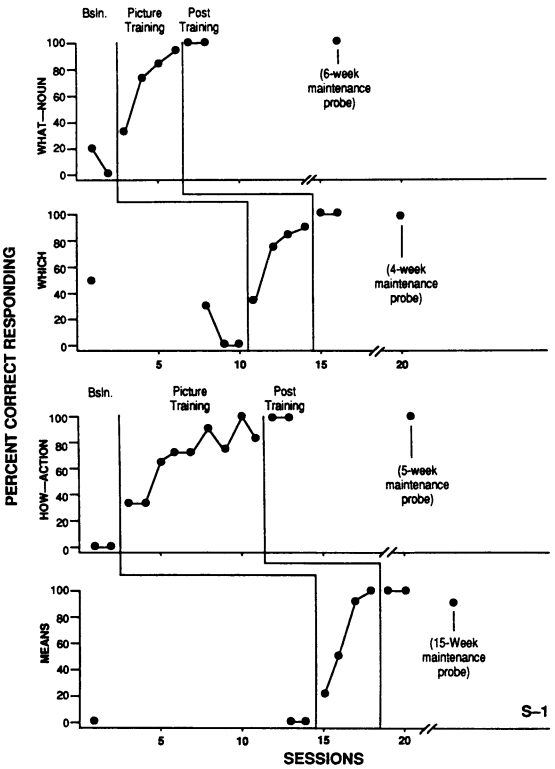

Figure 1. Percentage of correct responding on training stimuli for Student 1.

RESULTS

Training Stimuli

Figures 1 through 4 display the percentage of correct responses during baseline probes on the training stimuli. All of the students exhibited low levels of correct responding across question forms and subcomponents (M = 7.9%; range, 0% to 70%). The picture-training strategy was effective in increasing rates of correct responding to wh— questions when magazine pictures were the referents. The students required between three and nine training sessions to master each subcomponent (M = 5.18). All students, after meeting training criteria, successfully met the same criteria on posttraining probes administered by persons who had not been involved in the training. Maintenance probe data collected across a range of 1 to 68 weeks

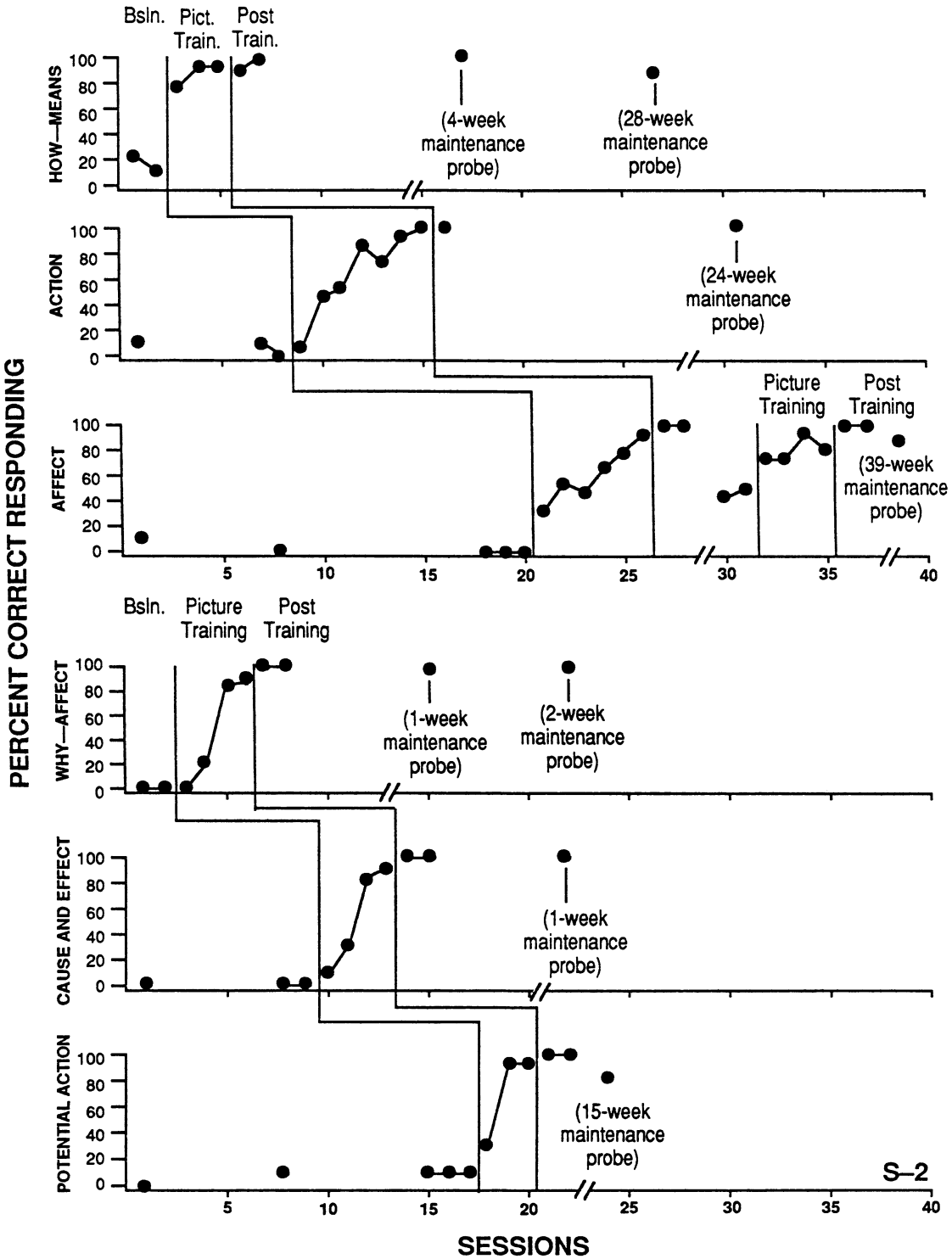

Figure 2. Percentage of correct responding on training stimuli for Student 2.

(M = 12) indicate that high levels of correct responding to pictures maintained without additional training.

Generalization

Results of probes to storybook and natural-context questions are displayed in Figures 5 through 8. Correct responding prior to training was very low for all students on both storybook and natural-context probes (M = 9.21%; range, 0% to 37%). Results of generalization probes following training are also displayed in Figures 5 through 8. Students responded at the criterion level on 22 of 34 (65%) subcomponents immediately following picture training. On the remaining 12 subcomponents, all students demonstrated higher rates of correct responding than in baseline, though not sufficient to meet criterion. For these items, the students were provided with specific generalization programming. The type (interspersal or direct) and focus (storybook or natural context) varied across students and subcomponents, as determined by individual probe results. Two to five (M = 4.5) remedial interspersal training sessions were necessary before students met criterion. Four direct story training sessions were necessary for both students trained in this condition.

Results of maintenance probes on natural-context questions were similar to those with pictures. Probes obtained across a range of 5 to 68 weeks showed that Students 1 through 3 maintained high levels of correct responding (M = 80%; range, 70% to 90%) on all subcomponents. Student 4 continued to respond correctly on two of three subcomponents (M = 85%; range, 80% to 90%); however, correct responding to the "why" subcomponent affect (24 weeks) was 20%.

Results of spontaneous question-asking probes indicate that no student ever asked a wh— question under any experimental condition.

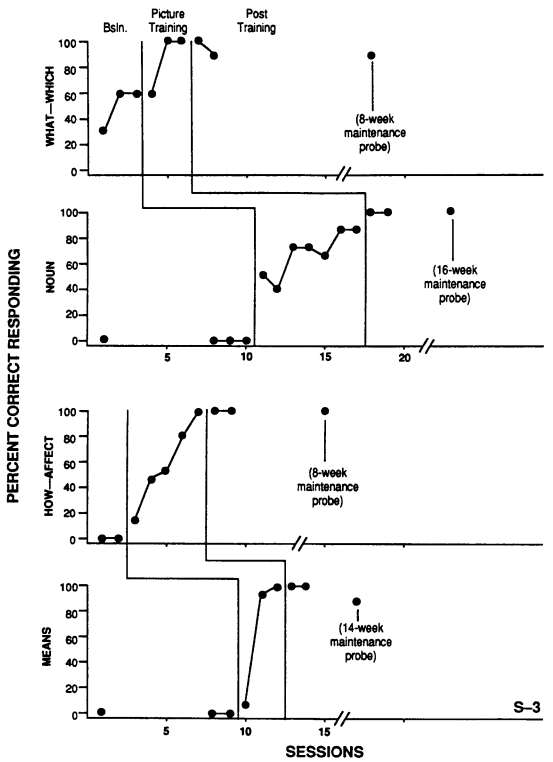

Figure 3. Percentage of correct responding on training stimuli for Student 3.

DISCUSSION

The results of the present study can be summarized as follows: All students demonstrated fairly rapid acquisition of targeted question forms, thus replicating the results of the Krantz et al. (1981) investigation. Additionally, all students performed at or above criterion on posttraining probes administered by persons not involved in training. Furthermore, maintenance probe results indicate that, for all but one student/subcomponent, responses to training stimuli and natural-context questions maintained over time.

Following picture training, all students demonstrated some increases in correct responding to generalization items; however, students failed to meet generalization criteria on 35% of storybook and natural-context stimuli. The interspersal condition was effective in promoting acceptable levels of generalization to 10 sets of previously unmastered stimuli, with the remaining two sets requiring direct training.

Several variables related to both students and instructional procedures may help to explain the

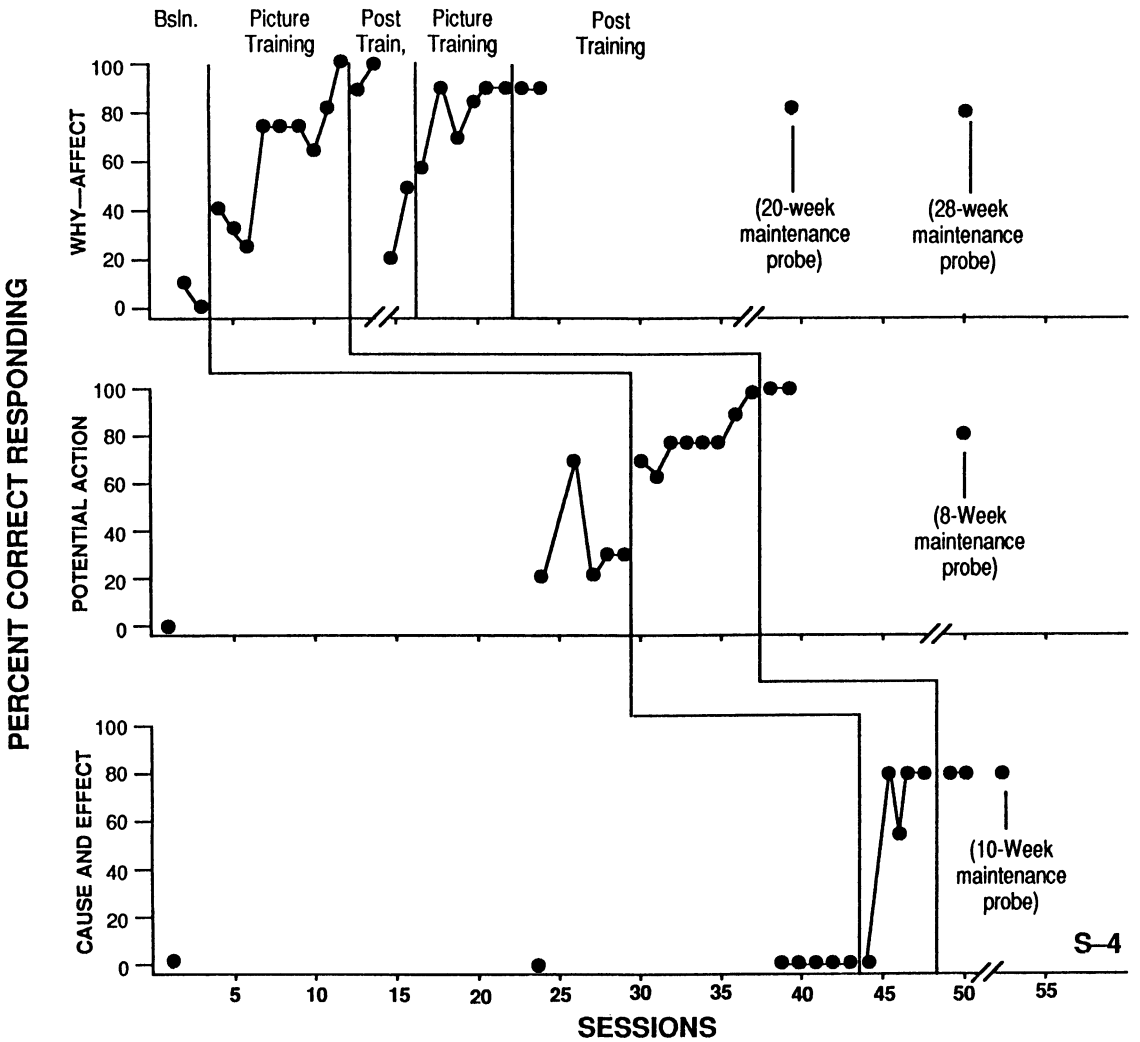

Figure 4. Percentage of correct responding on training stimuli for Student 4.

generalization results. A primary area of analysis relevant to the interpretation of generalization results concerns the nature of the cues that should control responding to the various subcomponents and the degree to which such cues can be presented within the context of picture training. For any given question, there exists one or more relevant cues that the respondent must attend to in order to formulate a correct response. These cues vary in number and complexity across question forms and subcomponents. Analysis of each of the subcomponents targeted in this study indicates that responses to half of these ("what," both subcomponents; "how," means and action) rely on relevant cues that are visible to the student. For example, when a student was asked, "How do you brush your teeth?", the relevant cue (i.e., the toothbrush) was plainly visible. Responding to the remaining subcomponents ("why," all subcomponents; "how," affect) requires attention to referents that are not all visible to the student. For example, the question, "Why is Johnny drinking?" could not be answered correctly by simply attending to the picture of a person drinking. The student had to be able to state the

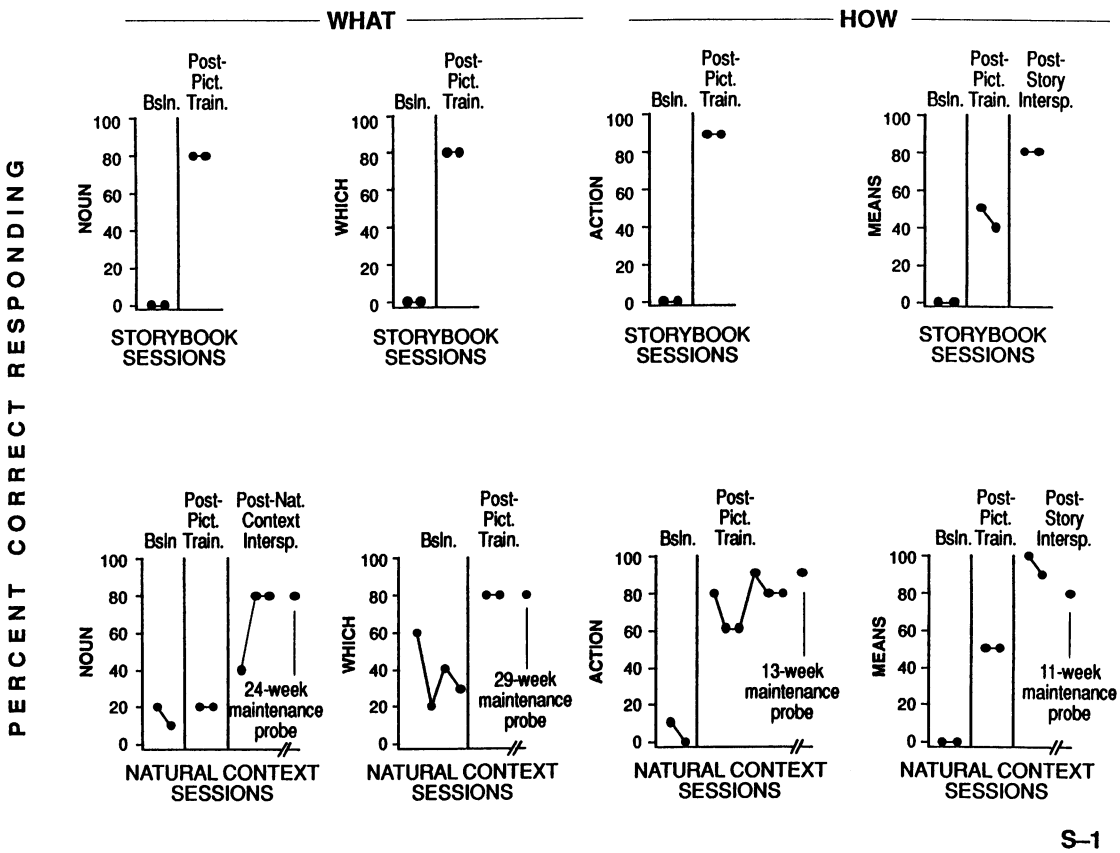

Figure 5. Percentage of correct responding during generalization probes for Student 1.

relationship between the act of drinking and the presumed motivation for that act (i.e., thirst) in response to a limited number of visible cues.

Comparison of the training data for the two classes (i.e., cues visible vs. not visible) of subcomponents indicates little difference between the average number of training trials to criterion. There was, however, a difference in the results for the generalization questions; the criterion was met on 83% of probes to questions where all relevant stimuli were visible but was met on only 42% of those for which a correct response relied on cues that were not all visible to the student.

These findings may be in part clarified by the results of two studies of responding to wh— questions by nonhandicapped children. Parnell, Amerman, Patterson, and Harding (cited in Parnell & Amerman, 1984) examined effects in the type and availability of referential sources on correct responding to a variety of wh— questions. One of the clear findings was that questions with no visible referent produced the largest number of incorrect responses. Additionally, an analysis of responding to wh— questions across children of a broad age range indicates that "how" and "why" questions, particularly those for which there is no visible referent, are the most difficult for children to answer (Parnell & Amerman, 1984).

Although these studies were conducted with linguistically normal children and did not examine generalization, they seem to have bearing on the present results. The modeling and reinforcement strategies used in picture training were equally effective in teaching responses to training stimuli for all classes of questions; however, their effectiveness in teaching a generalized response appears to di-

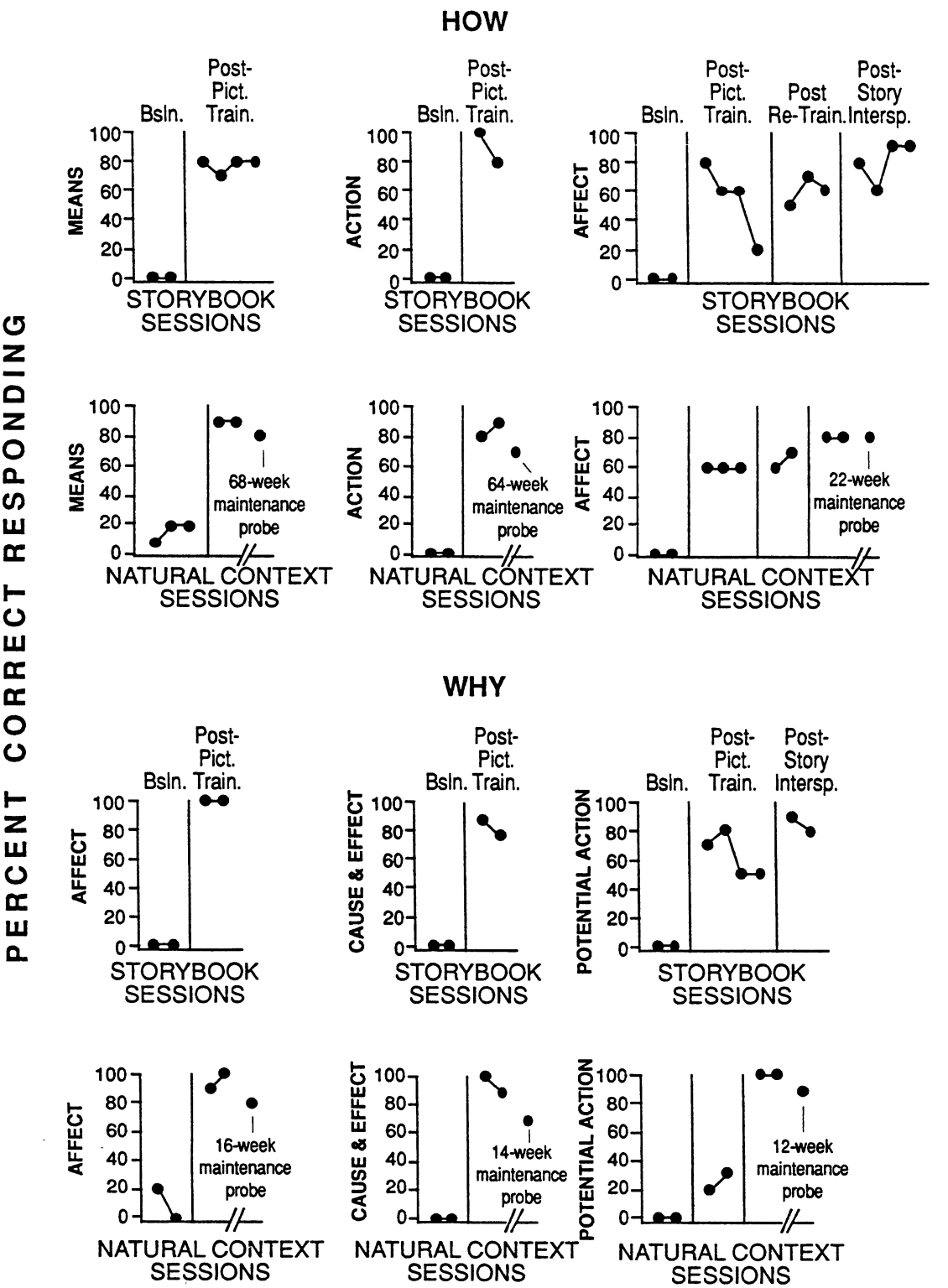

Figure 6. Percentage of correct responding during generalization probes for Student 2.

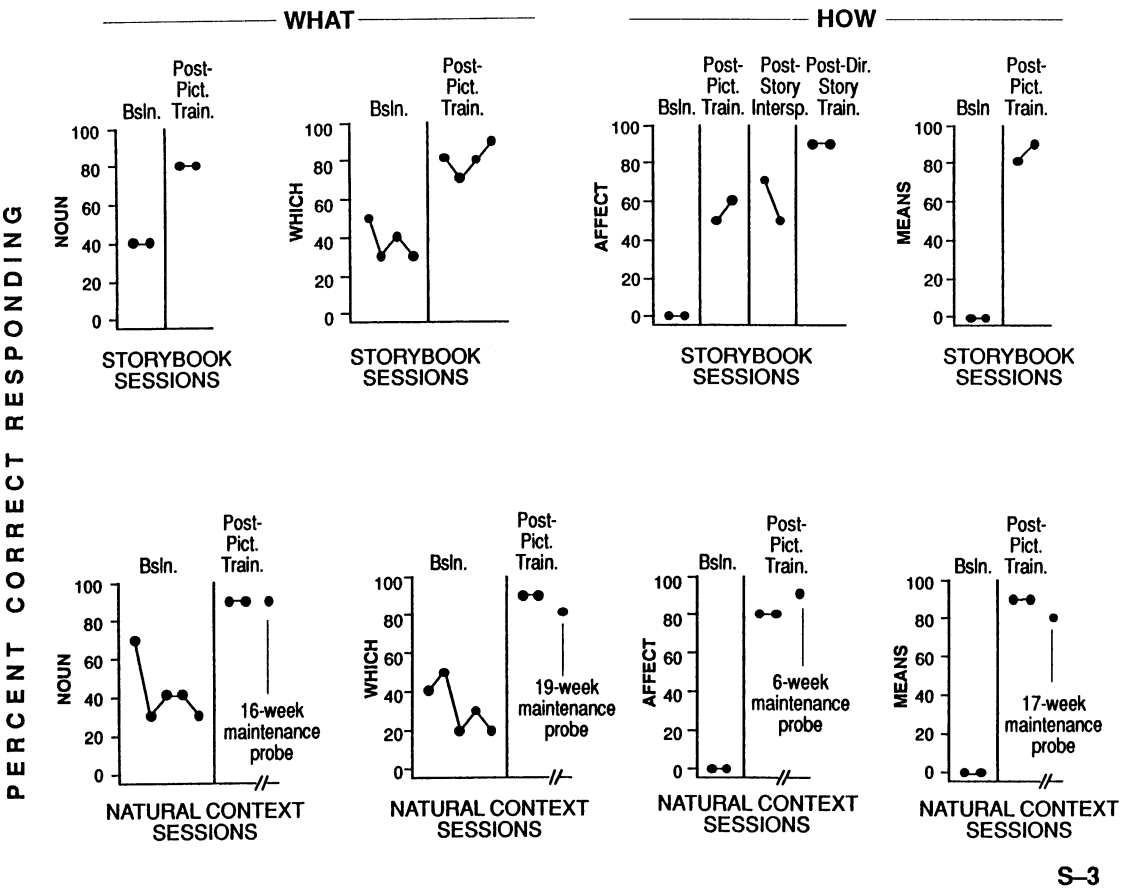

Figure 7. Percentage of correct responding during generalization probes for Student 3.

minish with more difficult and complex question forms. Generalization was higher for questions in which all relevant cues were visible, and generalization to "what" questions (87.5%) was higher than that to "how" (75%) and "why" (50%) questions.

Results of maintenance probes collected across a range of 5 to 68 weeks indicate that, once acquired, responding to training stimuli and natural-context questions maintained at 70% or better. The one exception was Student 4, who responded correctly to only 20% of "why"-affect questions. There are several possible explanations for these results. This student had just returned from a 6-week vacation and, although her responding to other subcomponents maintained at or above criterion, the time elapsed between maintenance and mastery probes was twice as long (24 weeks) for this subcomponent as for the other two (11 and 8 weeks). In addition, this subcomponent was clearly the most difficult for Student 4 to acquire and may also be one of the least frequently asked in a social context. Thus, these results suggest that maintenance activities may need to occur on a regular basis to ensure continued rates of correct responding on some question forms/subcomponents.

Interspersal training was an effective addition to the intervention package. One reason for its effectiveness may have been the broader range of relevant cues that could be presented. The interspersal conditions each provided five additional, novel examples of the subcomponent being trained. Furthermore, by interspersing storybook and/or natural-context questions with those from the picture stimuli, relevant cues were sampled that varied

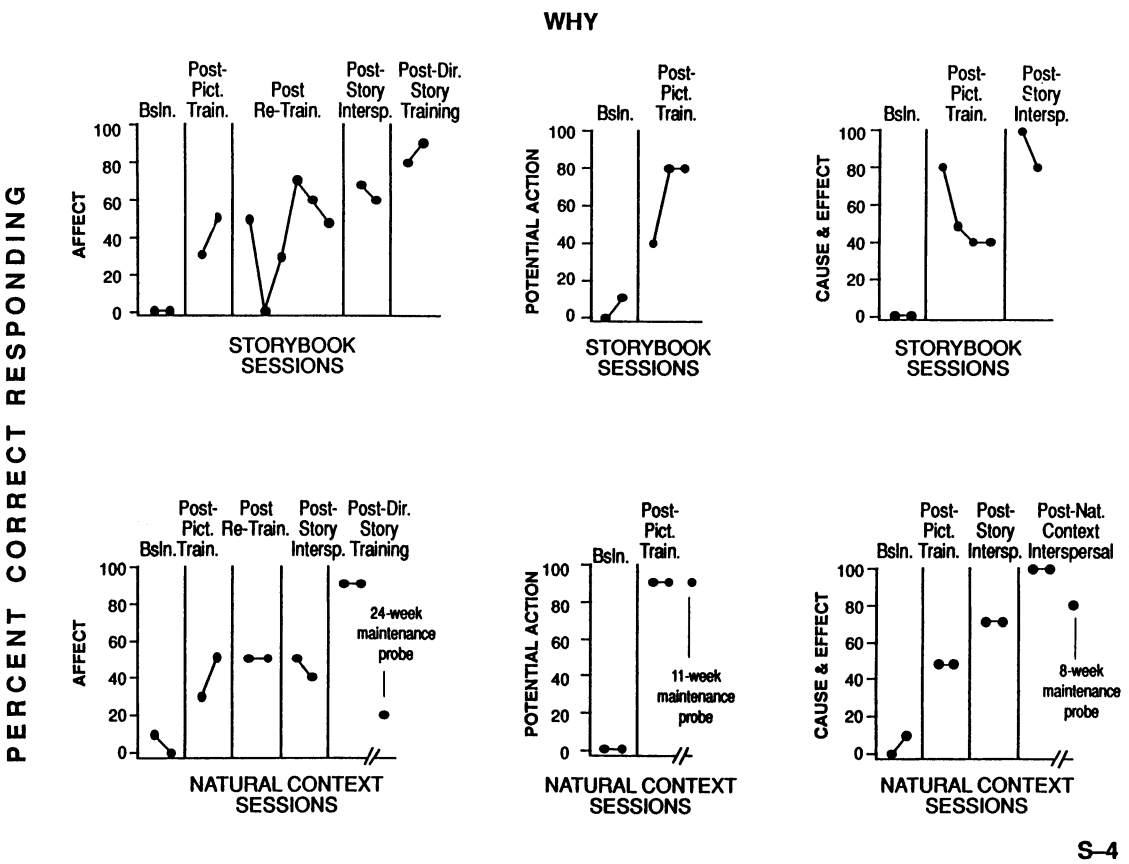

Figure 8. Percentage of correct responding during generalization probes for Student 4.

substantially from those presented in the picture-training condition alone. For example, tone of voice and facial expression are both relevant variables in answering a question about affect. In picture training, one of these variables (tone of voice) is not available to the learner; consequently, the response comes under the control of facial expressions alone. Such a restricted response strategy was not always sufficient to formulate a correct response when the students were required to make fine discriminations (e.g., between sad and angry) under more natural conditions. As such, the interspersal training may have served to sample more completely crucial cues in the environment that were not present in pictured situations.

The possibility of this variable contributing to the success of the interspersal strategy is supported by the data. On five different subcomponents, at least 1 student failed to meet generalization criteria on both storybook and natural-context questions. However, for three of the five, interspersal training with one type of question was sufficient to promote high levels of generalization to both types on subsequent probes. Thus, it is possible that providing additional training opportunities that more sufficiently sample relevant cues aided in teaching students a generalized response strategy.

Finally, several authors have demonstrated that an interspersal strategy can serve to increase motivation and attention by providing additional reinforcement opportunities (e.g., Dunlap & Koegel, 1980; Egel, Shafer, & Neef, 1984; Neef, Iwata, & Page, 1977, 1980). The interspersal procedure used in the present study incorporated both interspersal of known and novel items and presentation of a systematically varied set of training stimuli and

may have thus served to facilitate correct responding.

In summary, the picture-training strategy was effective in teaching a generalized response to wh— questions, especially for those in which relevant cues are visible. It appears, however, that a procedure such as the interspersal used in this study may be more effective in teaching children to respond to question forms that require responding on the basis of cues that are not visible. This may be particularly true for children who already tend to formulate a response on the basis of a restricted range of stimuli.

REFERENCES

Alpert, C., & Rogers-Warren, A. K. (1985). Communication in autistic persons: Characteristics and intervention. In S. Warren & A. K. Rogers-Warren (Eds.), *Teaching functional language: Generalization and maintenance of language skills* (pp. 123–155). Baltimore: University Park Press.

Bloom, L., & Lahey, M. (1978). *Language development and language disorders.* New York: John Wiley & Sons.

Brown, R. (1968). The development of wh— questions in child speech. *Journal of Verbal Learning and Verbal Behavior,* **7,** 279–290.

Carr, E. (1982). Sign language. In R. Koegel, A. Rincover, & A. Egel (Eds.), *Educating and understanding autistic children* (pp. 142–157). San Diego, CA: College Hill Press.

Carr, E. (1986). Behavioral approaches to language and communication. In E. Schopler & G. Mesibov (Eds.), *Communication problems in autism* (pp. 37–57). New York: Plenum Press.

Dunlap, G., & Koegel, R. (1980). Motivating autistic children through stimulus variation. *Journal of Applied Behavior Analysis,* **13,** 618–627.

Egel, A. L., Shafer, M. S., & Neef, N. A. (1984). Receptive acquisition and generalization of prepositional responding in autistic children: A comparison of two procedures. *Analysis and Intervention in Developmental Disabilities,* **4,** 285–298.

Engelmann, S., & Osborn, J. (1976). *DISTAR Language I: An instructional system.* Chicago: Science Research Associates.

Fay, W. H., & Schuler, A. L. (1980). *Emerging language in autistic children.* Baltimore: University Park Press.

Guess, D. (1969). A functional analysis of receptive language and productive speech: Acquisition of the plural morpheme. *Journal of Applied Behavior Analysis,* **2,** 55–64.

Guess, D., Sailor, W., Rutherford, G., & Baer, D. (1968). An experimental analysis of linguistic development: The productive use of the plural morpheme. *Journal of Applied Behavior Analysis,* **1,** 297–306.

Halle, J. (1982). Teaching functional language to the handicapped: An integrated model of natural environment teaching techniques. *Journal of the Association for Persons With Severe Handicaps,* **7,** 29–37.

Hood, L. A. (1977). *A longitudinal study of the development of the expression of causal relations in complex sentences.* Unpublished doctoral dissertation, Columbia University.

Horner, R. D., & Baer, D. M. (1978). Multiple probe technique: A variation on the multiple-baseline. *Journal of Applied Behavior Analysis,* **11,** 189–196.

Krantz, P., Zalenski, S., Hall, L., Fenske, E., & McClannahan, L. (1981). Teaching complex language to autistic children. *Analysis and Intervention in Developmental Disabilities,* **1,** 259–297.

Lord, C. (1986). Contribution of behavioral approaches to the language and communication of persons with autism. In E. Schopler & G. Mesibov (Eds.), *Communication problems in autism* (pp. 59–68). New York: Plenum Press.

Lovaas, O. I. (1977). *The autistic child: Language development through behavior modification.* New York: Irving Publishers.

Lutzker, J. R., & Sherman, J. A. (1974). Producing generative sentence usage by imitation and reinforcement procedures. *Journal of Applied Behavior Analysis,* **7,** 447–460.

Neef, N. A., Iwata, B. A., & Page, T. J. (1977). The effects of known item interspersal on acquisition and retention of spelling and sightwords. *Journal of Applied Behavior Analysis,* **10,** 738.

Neef, N. A., Iwata, B. A., & Page, T. J. (1980). The effects of interspersal training versus high-density reinforcement on spelling acquisition. *Journal of Applied Behavior Analysis,* **13,** 153–158.

Neef, N. A., Walters, J., & Egel, A. L. (1984). Establishing generative yes/no responses in developmentally disabled children. *Journal of Applied Behavior Analysis,* **17,** 453–460.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, Serial 149,* **38,** 1–2.

Parnell, M., & Amerman, J. (1984). Answers to wh— questions: Research and application. In T. Gallagher & C. A. Prutting (Eds.), *Pragmatic assessment and intervention issues in language.* San Diego, CA: College Hill Press.

Paul, L. (1985). Programming peer support for functional language. In S. F. Warren & A. K. Rogers-Warren (Eds.), *Teaching functional language: Generalization and maintenance of language skills* (pp. 289–307). Baltimore: University Park Press.

Ritvo, E., & Freeman, B. J. (1978). National Society for Autistic Children definition of the syndrome of autism. *Journal of Autism and Childhood Schizophrenia,* **8,** 162–167.

Sailor, W. (1971). Reinforcement and generalization of productive plural allomorphs in two retarded children. *Journal of Applied Behavior Analysis,* **4,** 305–310.

Schopler, E., & Mesibov, G. B. (1986). Introduction to communication problems in autism. In E. Schopler & G. B. Mesibov (Eds.), *Communication problems in autism* (pp. 3–13). New York: Plenum Press.

Schreibman, L., & Carr, E. G. (1978). Elimination of echolalic responding to questions through the training of a generalized verbal response. *Journal of Applied Behavior Analysis,* **11,** 453–463.

Schuler, A. L. (1980). Teaching functional language. In B. Wilcox & A. Thompson (Eds.), *Critical issues in the education of autistic children* (pp. 154–178). Washington, DC: U.S. Department of Education, Office of Special Education.

Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis,* **10,** 349–368.

Twardosz, S., & Baer, D. M. (1973). Training two severely retarded adolescents to ask questions. *Journal of Applied Behavior Analysis,* **6,** 655–661.

Wetherby, A. M. (1986). Ontogeny of communicative functions in autism. *Journal of Autism and Developmental Disorders,* **16,** 295–316.

Wolf, M. M., Risley, T., & Mees, H. (1964). Application of operant conditioning procedures to the behavior problems of an autistic child. *Behaviour Research and Therapy,* **1,** 305–312.

Received June 8, 1987
Initial editorial decision September 9, 1987
Revisions received January 10, 1988; June 18, 1988
Final acceptance January 30, 1989
Action Editor, Ron Van Houten